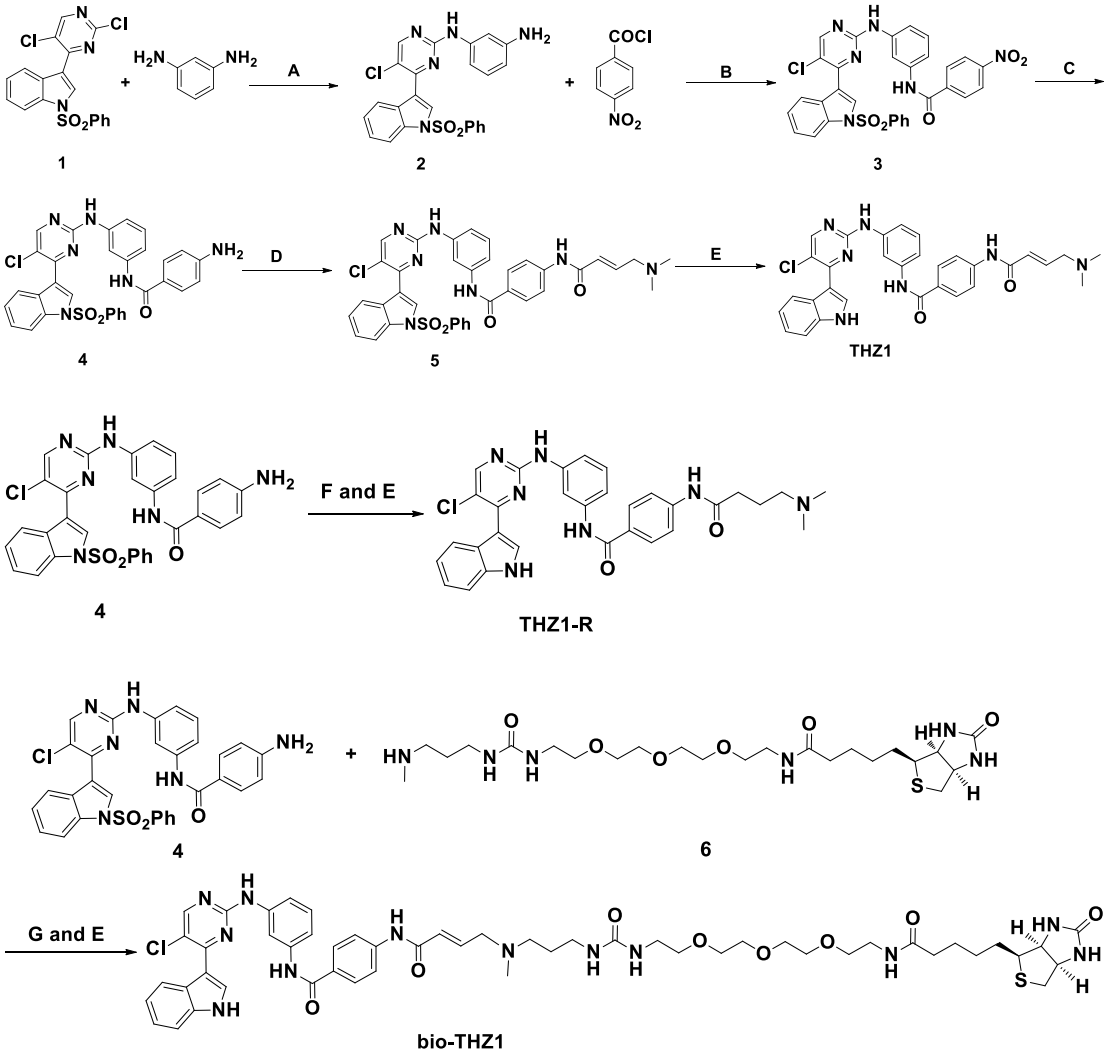
Supplementary Methods:

Chemistry.



Reagents and conditions: (A) NMP, 130°C (B) pyridine, 80 °C (C) $SnCl_2$, 80°C, ethyl acetate and methanol (D)*(E)*-4-bromobut-2-enoyl chloride, acetonitrile and then $NHMe_2$, 0°C-RT (E) 1M NaOH, 1,4-dioxane, RT (F) 4-chlorobutanoyl chloride, $NHMe_2$ (G) *(E)*-4-bromobut-2-enoyl chloride, acetonitrile and then biotin linker 6, 50°C, THF.

Chemistry. All solvents and reagents were used as obtained. 1H NMR spectra were

recorded with a Varian Inova 600 NMR spectrometer and referenced to dimethylsulfoxide. Chemical shifts are expressed in ppm. Mass spectra were measured with Waters Micromass ZQ using an ESI source coupled to a Waters 2525 HPLC system operating in reverse mode with a Waters Sunfire C18 5 μm, 4.6 mm x 50 mm column. Purification of compounds was performed with either a Teledyne ISCO CombiFlash Rf system or a Waters Micromass ZQ preparative system. The purity was analyzed on an above-mentioned Waters LC-MS Symmetry (C18 column,4.6 mm x 50 mm, 5 μM) using a gradient of 5-95% methanol in water containing 0.05% trifluoroacetic acid (TFA). Detailed synthetic schemes and characterization data are presented in the supplementary data.

N1-(5-chloro-4-(1-(phenylsulfonyl)-1H-indol-3-yl)pyrimidin-2-yl)benzene-1,3-diamine (2) 3-(2,5-dichloropyrimidine-4-yl)-1-1-(phenylsulfonyl)-1H-indole (403 mg, 1.0 mmol) and benzene-1,3-diamine (216 mg, 2.0 equiv) were dissolved in NMP (5 mL). After heating at 130°C for 3 hrs, the solution was cooled down to room temperature and then was diluted with ethyl acetate (100 mL). The resulted solution was washed with sat. $NaHCO_3$, water and brine. After drying with $MgSO_4$, the solvent was removed and the product was obtained by flash chromatography with dichloromethane/methanol (10/1) as eluent. (350 mg, yield 73%) LC-MS: *m/z* (M+H) 476; ^{1}H NMR (400 MHz, DMSO-*d6*) 9.53 (s, 1H), 8.61 (s, 1H), 8.57 (m, 1H), 8.37 (t, d = 7.6 Hz, 1H), 8.08 (m, 2H), 8.00 (d, *J* = 7.6 Hz, 1H), 7.71 (t, *J* = 7.2 Hz, 1H), 7.51 (d, *J* = 8.0 Hz, 2H), 7.43 (t, *J* = 7.2 Hz, 1H), 7.33 (t, *J* = 7.2 Hz, 1H), 6.87 (m, 3H), 6.22 (m, 1H), 4.93 (s, 2H)

N1-(3-(5-chloro-4-(1-(phenylsulfonyl)amino)-1H-indol-3-yl)pyrimidin-2-

yl)amino)phenyl)-4-nitrobenzamide (3) To a pyridine solution of free amine compound 2 (47.5 mg, 0.1 mmol) was added benzoyl chloride (22.0 mg,1.2 equiv). After stirring for 2 hrs at 80°C, the reaction mixture was concentrated and the crude was purified by HPLC to give pure product as a TFA salt. (52 mg, 83%) LC-MS: *m/z* (M+H) 625; ^{1}H NMR(400 MHz, DMSO) 10.51 (s, 1H), 9.93 (s, 1H), 8.65 (d, *J* = 6.4 Hz, 2H), 8.34 (m, 3H), 8.10 (m, 5H), 7.98 (d, *J* = 8.0 Hz, 1H), 7.12 (t, *J* = 7.2 Hz, 1H), 7.61 (t, *J* = 8.0 Hz, 2H), 7.50 (d, *J* = 9.2 Hz, 1H),7.37 (m, 2H), 7.25 (m, 2H).

4-amino-N-(3-(5-chloro-4-(1-(phenylsulfonyl)amino)-1H-indol-3-yl)pyrimidin-2-yl)amino)phenyl)benzamide (4) The nitro compound 3 (52 mg, 0.083 mmol) was suspended in ethyl acetate/methanol (5:1, vol/vol, 10 mL) and the resulted suspension was treated with $SnCl_2$ (40 mg, 2.5 equiv). After stirring for 2 hrs at 80°C, the reaction mixture was cooled down to room temperature and then was poured into sat. $NaHCO_3$ solution (10 mL). The mixture was stirred for 10 minutes and then was extracted with chloroform/2-propanol (4:1, vol/vol, 50 mL). The organic layer was washed with water and brine, dried over $MgSO_4$, filtered through a pad of celite and concentrated under reduced pressure. The crude was purified by HPLC to provide the compound 4. (32 mg, 65%) LC-MS: *m/z* (M+H) 595; ^{1}H NMR(600 MHz, DMSO) 9.82 (s, 1H), 9.71 (s, 1H), 8.64 (s, 1H), 8.62 (s, 1H), 8.36 (d, *J* = 8.4 Hz, 1H), 8.08 (d, *J* = 8.4 Hz, 2H), 8.06 (d, *J* = 7.8 Hz, 1H), 7.97 (d, *J* = 7.8 Hz, 1H), 7.72 (t, *J* = 7.2 Hz, 1H), 7.67 (d, *J* = 8.4 Hz, 2H), 7.60 (d, *J* = 7.8 Hz, 2H), 7.38 (m, 2H), 7.33 (d, *J* = 7.8 Hz, 1H), 7.25 (t, *J* = 7.8 Hz, 1H), 7.20 (t, *J* = 7.8 Hz, 1H), 6.60 (d, *J* = 8.4 Hz, 2H).

(E)-N-(3-((5-chloro-4-(1-(phenylsulfonyl)-1H-indol-3-yl)pyrimidin-2-yl)amino)phenyl)-4-(4-(dimethylamino)but-2-enamido)benzamide (5) To a solution of free amino compound 4 (60 mg, 0.11 mmol) in acetonitrile (5 mL) was added *N*,*N*-Diisopropylethylamine (40 uL) and (*E*)-4-bromobut-2-enoyl chloride (40 mg, 2.0 equiv) in dichloromethane (1 mL) at 0°C dropwise. After stirring for 5 minutes, dimethylamine (1M in THF, 2 mL) was added and the solution was allowed to stir at room temperature for 2 hrs. The solvent was then removed and the crude was purified by HPLC to give the product (58 mg, 82%). LC-MS: *m/z* (M+H) 706; ^{1}H NMR(600 MHz, DMSO) 10.65 (s, 1H), 10.09 (s, 1H), 9.87 (s, 1H), 8.64 (s, 1H), 8.62 (s, 1H), 8.36 (d, *J* = 8.4 Hz, 1H), 8.08 (m, 3H), 7.97 (d, *J* = 7.8 Hz, 1H), 7.89 (d, *J* = 8.4 Hz, 2H), 7.79 (d, *J* = 8.4 Hz, 2H), 7.70 (t, *J* = 7.8 Hz, 1H), 7.62 (t, *J* = 7.8 Hz, 2H), 7.45 (d, *J* = 7.8 Hz, 1H), 7.36 (m, 2H), 7.24 (m, 2H), 6.80 (m, 1H) 6.46 (d, *J* = 15.0 Hz, 1H), 3. 71 (s, 2H), 2.61 (s, 6H).

(E)-N-(3-((5-chloro-4-(1H-indol-3-yl)pyrimidin-2-yl)amino)phenyl)-4-(4-(dimethylamino)but-enamido)benzamide (THZ1) To a solution of compound 5 (58 mg, 0.08 mmol) in 1,4-dioxane (2 mL) was added 1.0 M NaOH (2 mL). The solution was stirred at room temperature for 2 hrs and then was quenched with 1.0 M HCl (2 mL). The solution was extracted with chloroform/2-propanol (4/1, vol/vol, 20 mL) and the organic layer was washed with water, brine and dried with $MgSO_4$. The solvent was removed under reduced pressure and the crude was purified by HPLC to provide THZ1 as TFA salt. (33 mg, 72%) LC-MS: *m/z* (M+H) 566; ^{1}H NMR (600 MHz, DMSO-*d6*) 11.91 (s, 1H), 10.59 (s, 1H), 10.09 (s, 1H), 9.65 (s, 1H), 8.62 (d, *J* = 7.8 Hz, 1H), 8.51 (d, *J* = 3.6

Hz, 1H), 8.44 (s, 1H), 8.14 (s, 1H), 7.94 (d, *J* = 8.4 Hz, 2H), 7.77 (d, *J* = 8.4 Hz, 2H), 7.50 (d, *J* = 7.8 Hz, 1H), 7.47 (d, *J* = 7.8 Hz, 1H), 7.37 (d, *J* = 7.8 Hz, 1H), 7.26 (t, *J* = 7.8 Hz, 1H), 7.15 (t, *J* = 7.8 Hz, 1H), 7.07 (t, *J* = 7.8 Hz, 1H), 6.77 (m, 1H), 6.49 (d, *J* = 15.0 Hz, 1H), 3.95 (d, *J*= 6.0 Hz, 2H), 2.79 (s, 6H); ^{13}C NMR (400 MHz, DMSO-d_6) 165.2, 163.9, 158.6, 158.5, 157.5, 142.5, 142.0, 140.9, 139.7, 136.5, 131.6, 129.9, 129.1 (2C), 128.8, 126.5, 126.4, 123.6, 122.8, 121.1, 118.8 (2C), 115.9, 115.6, 115.1, 113.1, 112.2, 111.3, 59.9, 45.3.

N-(3-((5-chloro-4-(1H-indol-3-yl)pyrimidin-2-yl)amino)phenyl)-4-(4-(dimethylamino)butanamido)benzamide (THZ1-R) To a solution of free amino compound 4 (30 mg, 0.05 mmol), *N,N*-diisopropylethylamine (20 uL) in acetonitrile (2 mL) was added 4-chlorobutanoyl chloride (15 mg, 2.0 equiv) in dichloromethane (1 mL). After stirring for 5 minutes at 0°C, dimethylamine (1M in THF, 2 mL) was added and the reaction mixture was heated to 50°C for 2 hrs. The solvent was removed and the crude was treated with 1,4-dioxane (2 mL) and 1 M NaOH solution (2 mL). The resulted solution was stirred at room temperature for 2 hrs. The reaction was quenched with 1M HCl (2 mL) and then was extracted with chloroform/2-propanol (4/1, vol/vol, 20 mL). The organic layer was washed with water, brine and dried with $MgSO_4$. The concentration under reduced pressure to give the crude, which was purified by HPLC to provide THZ1-R as TFA salt (17.0 mg, 60%) LC-MS: *m/z* (M+H) 568; ^{1}H NMR (600 MHz, DMSO-*d6*) 11.93 (s, 1H), 10.32 (s, 1H), 10.06 (s, 1H), 9.64 (s, 1H), 8.62 (d, *J* = 7.8 Hz, 1H), 8.51 (d, *J* = 3.6 Hz, 1H), 8.44 (s, 1H), 8.13 (s, 1H), 7.91 (d, *J* = 8.4 Hz, 2H), 7.71 (d, *J* = 8.4 Hz, 2H), 7.50 (d, *J* = 7.8 Hz, 1H), 7.47 (d, *J* = 7.8 Hz, 1H), 7.37 (d, *J* =

7.8 Hz, 1H), 7.24 (t, *J* = 7.8 Hz, 1H), 7.15(t, *J* = 7.8 Hz, 1H), 7.07 (t, *J* = 7.8 Hz, 1H), 3.09 (m, 2H), 2.78(s, 3H), 2.77 (s, 3H), 2.47(t, *J* = 7.2 Hz, 2H), 1.95 (m, 2H); ^{13}C NMR (400 MHz, DMSO-*d6*)170.0 165.2, 158.6, 158.5, 157.5, 142.4, 140.9, 139.7, 131.8, 129.7 (2C), 129.0, 128.8, 126.5, 123.6, 122.8, 121.1, 118.6 (2C), 116.0, 115.6, 115.1, 113.1, 112.2, 111.3, 56.6, 42.6, 33.2, 20.4.

(E)-N-(3-((5-chloro-4-(1H-indol-3-yl)pyrimidin-2-yl)amino)phenyl)-4-(5-methyl-19-oxo-23-(2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-9,12,15-trioxa-5,18-diazatricos-2-enamido)benzamide (bio-THZ1) The free amino compound (**4**) (30 mg, 0.05 mmol) was dissolved in acetonitrile (2.0 mL) and to the solution, *N,N*-diisopropylethylamine (20 uL) and (*E*)-4-bromobut-2-enoyl chloride (20 mg, 2.0 equiv) in dichloromethane (1 mL) were added at 0°C. After stirring for 5 minutes, the reaction was quenched with sat. $NaHCO_3$ and was extracted with chloroform/2-propanol (4/1, vol/vol, 20 mL). The organic layer was washed with water and brine. After drying and removal of the solvent, the crude was dissolved in THF (1 mL) and to the solution, the biotin-linker 6 (25 mg) was added. The solution was then heated to 50°C for 2 hrs. After removal of THF under reduced pressure, the crude was treated with 1,4-dioxane (2 mL) and 1M NaOH solution (2 mL) and then was stirred at room temperature for 2 hrs. After adding 1M HCl solution (2 mL), the reaction solution was extracted with chloroform/2-propanol (4/1, vol/vol, 20 mL) and the organic layer was washed with water, brine and dried over $MgSO_4$. Removing the solvent to give the crude, which was purified by HPLC to provide bio-THZ1 as TFA salt (16.0 mg, 30%) LC-MS: *m/z* (M+H) 1053, ^{1}H NMR (600 MHz, DMSO-*d*6) 11.93 (s, 1H), 10.67 (s, 1H), 10.17 (s, 1H), 9.90 (br, 1H), 9.73 (s,

1H), 8.70 (d, *J* = 7.8 Hz, 1H), 8.55 (d,*J* = 3.0 Hz, 1H), 8.52 (s, 1H), 8.21 (m, 1H), 8.02 (d, *J* = 8.0 Hz, 2H), 7.89 (t, *J* = 6.0 Hz, 1H), 7.85 (d, *J* = 8.0 Hz, 2H), 7.57 (d, *J* = 8.4 Hz, 1H), 7.54 (d, *J* = 8.4 Hz, 1H), 7.46 (t, *J* = 7.8 Hz, 1H), 7.32 (t, *J* = 8.4 Hz, 1H), 7.22 (t, *J* = 8.4 Hz, 1H), 7.15 (t, *J* = 8.4 Hz, 1H), 6.87 (m, 1H), 6.58 (d, *J* = 16.2 Hz, 1H), 6.48 (br, 1H), 6.27 (br, 1H), 6.15 (br, 1H), 4.35 (m, 2H), 4.17 (m, 2H), 4.10- 2.55(m, 26H) 2.12 (t, *J* = 7.2 Hz, 2H), 1.82 (m, 2 H), 1.65 (m, 1H), 1.56 (m, 3H), 1.34 (m, 2H).

Reagents and Antibodies. Flavopiridol and Actinomycin D were purchased from Sigma (cats# F3055 and A1410). The following antibodies were used for immunoblots: RNAPII CTD Ser-2 (cat# 04-1571), Ser-5 (cat# 04-1572), and Ser-7 (cat# 04-1570) phospho-antibodies (Millipore); Total RNAPII (Santa Cruz cat# sc-17798); CDK7 (Santa Cruz cat# sc-365075 (C-terminal) and sc-723 (N-terminal)); CDK9 (Bethyl cat# A303—493A); phospho-CDK9 (Cell Signaling cat# 2549); c-MYC (Epitomics cat# 1472-1); MCL-1 (Millipore cat# MABC43); BCL-2 (Cell Signaling cat# 2876); Survivin (Cell Signaling cat# 2803); BCL-XL (Bethyl cat# A300-284A); XIAP (Cell Signaling cat# 2042); CDK1 (Bethyl cat# A303-664A); CDK2 (Bethyl cat# A301-812A); CDK4 (Cell Signaling cat# 2906); CDK6 (Santa Cruz cat# sc-177); CDK8 (Santa Cruz cat# sc-1521); cyclin H (Bethyl cat# A301-674A); MAT1 (Santa Cruz cat# sc-13142); XPB (Santa Cruz cat# sc-293); XPD (Santa Cruz cat# sc-101174); p62 (Santa Cruz cat# sc-25329); cyclin K (Bethyl cat# A301-939A); TAL1 (Millipore cat# 04-123); GATA3 (Santa Cruz cat# sc-9009); RUNX1 (AML1) (Abcam cat# ab92336); MYB (Millipore cat# 05-175); phospho-CDK1 (Cell Signaling cat# 9114); phospho-CDK2 (Cell Signaling cat# 2561); PARP (Cell Signaling cat# 9542); Streptavidin-HRP (Thermo Scientific cat# N100); Flag

M2 (Sigma cat# F3165); and α-Tubulin DM1α (Sigma cat# T9026). The following ChIP-grade antibodies were used for ChIP-seq: RNAPII (Santa Cruz cat# sc-899); CDK7 (Abcam cat# ab9516); and H3K27Ac (Abcam cat# AB4729).

Cell Culture. Jurkat, Loucy, KOPTK1, and DND-41 T-ALL cell lines were grown in RPMI-1640 with 1% glutamine. Hela S3, HCT116 and 293A cells were grown in DMEM medium. Hela S3 expressing both pLVX-Tet-On Advanced and tetracycline –inducible FLAG-tagged CDK7 pLVX-Tight-Puro constructs were grown in DMEM medium supplemented with 1 mg/mL G418 sulfate and 2 μg/mL puromycin. Expression of CDK7 transgenes was induced with 2 μg/mL doxycycline for 24 hrs prior to compound treatment and for duration of experiment. HCT116 expressing FLAG-tagged CDK7 proteins were grown in DMEM medium supplemented with 2 μg/mL puromycin. 293A cells expressing FLAG-CDK12 were grown in medium supplemented with 1 mg/mL G418 sulfate. BJ Fibroblasts were grown in ATCC Eagle's Minimum Essential Medium. hTERT-RPE1 cells were grown in DMEM and Ham's F-12, 50/50 Mix (Cellgro – Fisher) supplemented with sodium bicarbonate (Invitrogen) and hygromycin B (Invitrogen) to final concentrations of 0.348% and 10 $\mu g \cdot mL^{-1}$ respectively. All cell lines were supplemented with 10% FBS (Sigma) and cultured at 37^{o}C in a humidified chamber in the presence of 5% CO_2, unless otherwise noted. Cell lines have not recently been authenticated, but had tested negative for mycoplasma.

Cloning, Mutagenesis, and Expression. hsCDK7 WT and D155A cDNA in pcDNA3 were purchased from Addgene (plasmid #s 14647 and 14648). C312S mutagenesis was

performed using the following primer GCTGCCAAGACCAAACTCTCCAGTGGAAACCTTAAAGG. For expression in HCT116 cells, N-terminally FLAG-tagged CDK7 variants were generated by sub-cloning the CDK7 cDNAs (above) into the FLAG-pLJM1 lentiviral expression vector[1] (provided by the laboratory of David Sabatini) using modified SalI/Not1 restriction sites. For expression in Hela S3 expressing cells, FLAG-tagged CDK7 variants were further sub-cloned into pLVX-Tight-Puro. FLAG-CDK7 constructs were co-transfected with the psPax2 envelope and CMV VSV-G packaging plasmids into actively growing HEK-293T using FuGENE 6 transfection reagent as described previously[1]. Virus-containing supernatants were collected at 48 hrs after transfection, filtered to eliminate cells, and HCT116 or Hela S3 pLVX-Tet-On Advanced targets cells were infected in the presence of 8 μg/ml polybrene. Forty-eight hrs later cells were selected with puromycin and passaged further.

Immunoblotting. Whole cell lysates for immunoblotting were prepared by pelleting cells from each cell line at 4^{o}C (1,200 rpm) for 5 min using a Sorvall Legend centrifuge (Thermo Fisher Scientific). The resulting cell pellets were washed 1X with ice-cold 1X PBS and then resuspended in lysis buffer containing 50 mM TrisHCl pH 8.0, 150 mM NaCl, 1% NP-40, 0.1% SDS, 5 mM EDTA with protease (Roche cat# 04 693 159 001) and phosphatase inhibitors (Roche cat# 04 906 837 001). Whole cell lysates were collected and snap-frozen in liquid nitrogen before being stored at -80^{o}C. Protein concentrations were determined by using the Biorad DC protein assay kit (Bio-Rad, #500-0111). Whole cell lysates were loaded into NOVEX 4%–20% Tris Mini Gel

(Invitrogen, Carlsbad, CA) and separated by electrophoreses at 150 V for 2 hr. The gels were then transferred onto nitrocellulose membrane (Biorad cat# 162-0115) and blocked by incubation with 5% dry milk in TBST (TBS with 0.2% Tween-20). Membranes were probed using antibodies raised against the indicated proteins. Chemiluminescent detection was performed with appropriate secondary antibodies and developed using high-resolution BioMax MR film (Kodak, category: 870 13012).

Inhibitor treatment experiments. Time-course experiments such as those described in Supplementary Fig. 4a were conducted to determine the minimal time required for full inactivation of CDK7. Here cells were treated with THZ1, THZ1-R, or DMSO for indicated amounts of time. Cells were lysed in lysis (as detailed above) buffer containing protease and phosphatase inhibitors and whole cell lysates were probed for the indicated proteins using standard immunoblotting techniques. For subsequent experiments cells were treated with compounds for 4 hrs as determined by time-course experiment described above, unless otherwise noted. For inhibitor washout experiments cells were treated with THZ1, THZ1-R, or DMSO for 4 hrs. Cells were then washed 3 times with fresh medium without inhibitors to effectively 'washout' the inhibitor. Fresh medium containing no inhibitors was then added back to the cells. Cells were then allowed to grow in medium without inhibitors for 0 to 6 hrs. 'N' signifies cells where medium was never washed out (*ie* – 10 hr incubation with compounds). Cells were then lysed at indicated time points and lysates were probed for RNAPII CTD phosphorylation and other specified proteins.

Pull down/ Immunoprecipitation (IP) experiments. Cells were treated with THZ1, THZ1-R, or DMSO for 4 hrs. Following treatment cells were washed 2-fold with cold PBS and then lysed in the following lysis buffer: 50 mM TrisHCl pH 8.0, 150 mM NaCl, 1% NP-40, 5 mM EDTA, 1mM DTT, and protease/phosphatase cocktails. Following clearance, lysates were treated with bio-THZ1 for pulldown (or CDK7 antibody pre-conjugated to Protein G agarose for IP) overnight at 4 ^{o}C. In the case of bio-THZ1 pulldown lysates were further incubated at room temperature for 3 hrs to increase the efficiency of covalent bond formation. Lysates were then incubated with streptavidin agarose for pulldown for an additional 2-3 hrs at 4^{o}C. Agarose beads were washed 5 times with lysis buffer (or with lysis buffer containing de-naturing 4M urea) and then boiled in 2X SDS for 10 minutes at 95 ^{o}C. SDS-page resolved precipitated proteins were probed for the indicated proteins. For CDK7 IPs, Protein G Agarose beads were washed 5 times with lysis buffer and the immunoprecipitated proteins were probed using the indicated antibodies.

Lanthascreen *In vitro* Kinase Assays. LanthaScreen® Eu Kinase Binding assays were conducted at Life Technologies in a time-dependent manner (20, 60, and 180 min.) using the manufacturer's specifications for each kinase indicated.

Radiometric *in vitro* kinase assays. For kinase assays following immunoprecipitation of FLAG-CDK7 protein from HCT116 or FLAG-CDK12 from 293A cellular lysates, cells were first treated with THZ1, THZ1-R, or DMSO for 4 hrs at 37^{o}C. Cells were then harvested by lysis in 50 mM TrisHCl pH 8.0, 150 mM NaCl, 1% NP-40, 5 mM EDTA,

and protease/phosphatase cocktails. Exogenous CDK7 or CDK12 proteins were immunoprecipitated from cellular lysates using FLAG antibody- conjugated agarose beads (Sigma cat# A2220). Precipitated proteins were washed with lysis buffer 6 times, followed by 2 washes with kinase buffer (40 mM Hepes pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 5% glycerol) and subjected to *in vitro* kinase assays at 30 ^oC for 45 minutes using 1 μg of the large subunit of RNAPII (RPB1) as substrate and 25 μM ATP and 10 μCi of ^{32}P ATP. Note: no additional inhibitors were added to the kinase reaction mixture, therefore any inhibition results from the activity of the compounds that were added directly to cells (*ie* - intracellular inhibition). This suggests that the inhibitors are either covalently tethered to the kinase or have strong non-covalent character. Kinase assays using recombinant CDK7/TFIIH/MAT1 (CAK complex, Millipore cat# 14-476) were conducted in the manner as described above using 25 ng of CAK complex per reaction. For kinase assays designed to test time-dependent inactivation of CDK7 kinase activity (Supplementary Fig. 1), CAK complex was pre-incubated with indicated concentrations of THZ1, THZ1-R, or DMSO in kinase buffer without ATP for 4 hrs at 30 ^oC prior to being subjected to kinase assay conditions as described above.

CDK7 C312S phenotypic rescue experiments. Heterogeneous populations of Hela S3 stably expressing either FLAG-CDK7 WT or C312S were enriched for single clones expressing high levels of CDK7 transgene compared to endogenous CDK7 protein levels. Clonal populations were expanded and maintained in 1 mg/ mL G418 sulfate and 2 μg/mL puromycin. Expression of FLAG-CDK7 transgenes was induced by the addition of 2 μg/mL doxycycline. Twenty-four hrs later THZ1 or DMSO was added at the

indicated concentrations and for the specified amount of time. Following compound treatment, cellular lysates were collected as previously described for immunoblots or mRNA was collected as detailed elsewhere for gene expression analysis.

Proliferation Assays. Proliferation assays were conducted using Cell Titer Glo assay kit (Promega cat# G7571). T-ALL cells grown in suspension were resuspended in fresh media containing THZ1, THZ1-R, DMSO, or other compounds at the indicated concentrations and then plated in 96-well plates at 20,000 cells/ well in a volume of 100 μL. Anti-proliferative effects of compounds were assessed following 72 hr incubations. For adherent cell lines, cells were seeded in 96-well plate at ~1,000-2,000 cells/ well. After 24 hrs cells were treated with inhibitors at the indicated concentrations and anti-proliferative effects were assessed after 72 hrs. To assess the effect of inhibitor washout on anti-proliferation of T-ALL suspension cell lines, cells were treated with THZ1, THZ1-R, or DMSO for 4 hrs. Inhibitor-containing medium was then removed as detailed above and cells were re-seeded in 96-well plates at a density of 20,000 cells/ well in 100 μL in medium without inhibitors for 72 hrs. Anti-proliferative effects were then determined using Cell Titer Glo as described in product manual by luminescence measurements on a Tecan Safire plate reader. All proliferation assays (excluding cancer cell line profiling) were performed in biological triplicate and error bars are +/- SD. IC_{50}s were determined using GraphPad Prism 6 non-linear regression curve fit.

Molecular Modeling.

A crystal structure (PDB code 1UA2)[2] of human CDK7 from the Protein Data Bank was adapted for the prediction of the covalent binding model of THZ1 to the C-terminal domain of CDK7. We modified the C-terminal domain of the CDK7 structure using a homology-modeling tool in Discovery Studio 3.5 (http://www.accelrys.com) because the crystal structure of human CDK7 contained only residues Glu13 to Asn311. THZ1 was docked onto a minimized crystal structure of the modified human CDK7 crystal structure using the covalent docking routine in Prime (Prime, version 3.1, Schrödinger, LLC, New York, NY, 2012). The covalent docking study was performed using default settings, keeping all residues fixed except for Cys312 modeled into the CDK7 crystal structure.

Nanoflow LC-MS/MS. Recombinant CAK complex (CAK complex, Millipore cat# 14-476) was incubated with THZ1 or DMSO for 4 hrs at 37 °C, and the reactions were resolved by SDS-PAGE. Bands corresponding to CDK7 were excised, and digested in gel with trypsin according to standard protocols. Extracted peptides were loaded via autosampler injection (NanoAcquity Sample Manager, Waters, Milford, MA) onto a precolumn (4 cm POROS 10R2, Applied Biosystems, Framingham, MA) and eluted with an HPLC gradient (NanoAcquity Binary Sample Manager, Waters; 0-35% B in 60 minutes; A=0.1 M acetic acid in water, B=0.1M acetic acid in acetonitrile). Peptides were resolved on a self-packed analytical column (12 cm Monitor C18, Column Engineering, Ontario, CA) and introduced to the mass spectrometer (LTQ Orbitrap XL or LTQ Orbitrap Velos) at a flow rate of ~30 nL/min (ESI spray voltage = 2.2 kV)[3]. The Orbitrap XL was operated in data dependent mode such that the 15 most abundant precursors were subjected to MS/MS (CAD with electron multiplier detection,

NCE=35%). To facilitate detection of peptides labeled with the inhibitor, acetonitrile was added to peptides (final concentration 25%) prior to injection. In addition, peptides were eluted using a modified HPLC gradient (10-70% B in 24 minutes) and analyzed on the Orbitrap Velos. Here, the mass spectrometer performed alternating CAD (EM detection, NCE=35%) and HCD MS/MS (NCE=35%, image current detection, resolution 7500 @ m/z 400) scans on the top 7 most abundant precursors and included additional targeted scans for various charge states predicted for the inhibitor labeled C312 peptide.

MS Data Analysis. The multiplierz desktop environment[4] was used to generate peak lists for search of MS/MS spectra against a database of CDK7 by Mascot 2.2.1. Precursor and product ion tolerances were 10 ppm and 0.6 Da, respectively. Search parameters included trypsin specificity, up to 2 missed cleavages, fixed carbamidomethylation (C, +57 Da), variable deamidation (NQ, -1 Da), oxidation (M, +16 Da), and inhibitor modification (CK, +565 Da). The doubly carbamidomethylated C312 containing peptide was quantified from MS scans using peak heights corresponding to triply and quadruply charged peptide precursors using 2 unlabeled CDK7 peptides for normalization.

Cancer cell line profiling.

Cell lines in high-throughput screening. All cell lines were sourced from commercial vendors. Cells were grown in RPMI or DMEM/F12 medium supplemented with 5% FBS and penicillin/streptavidin, and maintained at 37°C in a humidified atmosphere at 5% CO2. Cell lines were propagated in these two media in order to minimize the potential

effect of varying the media on sensitivity to therapeutic compounds in our assay, and to facilitate high-throughput screening. To exclude cross-contaminated or synonymous lines, a panel of 92 SNPs was profiled for each cell line (Sequenom, San Diego, CA) and a pair-wise comparison score calculated. In addition, we performed short tandem repeat (STR) analysis (AmpFlSTR Identifiler, Applied Biosystems, Carlsbad, CA) and matched this to an existing STR profile generated by the providing repository. More information on the cell lines screened, including their SNP and STR profiles is available on the Genomics of Drug Sensitivity in Cancer project website (www.cancerRxgene.org).

High-throughput cell line panel viability assay. Cells were seeded in 384-well microplates at ~15% confluency in medium with 5% FBS and penicillin/streptavidin. The optimal cell number for each cell line was determined to ensure that each was in growth phase at the end of the assay. Adherent cell lines were plated one day prior to treatment with a 9-point 2-fold dilution series of each compound using liquid handling robotics, and assayed at a 72-h time point. Suspension cell lines were treated with compound immediately following plating, incubated for 72 hrs. Cell viability was determined using resazurin. Quantitation of fluorescent signal intensity was performed using a fluorescent plate reader at excitation and emission wavelengths of 535/595 nM. All screening plates were subjected to stringent quality control measures: only plates with a coefficient of variation of less than 20% for control (DMSO treated) wells and a signal/noise ratio (blanks – no cells – wells versus DMSO treated wells) over 5 were included in data analysis. IC_{50}s for cell viability in the high-throughput screen were calculated as previously described[5].

Elastic net regression. Elastic net regression was performed as described in[5], with a 100 cross-validation using a random ten-fold partition of the samples. Genome features used for modeling were comprised of mRNA expression (13321 genes, U133A affymetrix array), copy number changes (360 genes from the cancer gene census on COSMIC http://cancer.sanger.ac.uk/cancergenome/projects/census/) and sequence variants or gene fusion across 83 cancer genes (see Supplementary Table 4 and [5]). For each of the 100 runs, a feature list is built for the drug comprised of genes, transcripts, and tissues with weights assigned to each. The final signature of markers for a drug consists of all features that appear in any of the 100 runs along with the statistics on the frequency that the feature appears and the average weight given to that feature over the 100 runs. The weights are used to assess effect sizes of features in the drug's marker signature. The effect size of a feature is calculated by multiplying the feature's weight by its standard deviation across the cell line panel. The effect size is therefore a normalization of the feature's weight to account for the different scales used to measure the different genomic features. Features with higher stability of correlation in cross-validation (f) are considered to have the highest confidence of truly being associated with drug response. The most significant features associated with drug response are those with both large frequency and effect size. For gene expression, a negative effect size corresponds to higher expression in sensitive cell lines. For mutation or copy number features (gene name followed by Mut or CP in Supplementary Table 5) negative effect indicates the presence of the mutation or abnormal copy number (amplification or deletion) in

sensitive cell lines, conversely positive effects indicate the presence or higher expression in more resistant cell lines.

Functional enrichment analysis using DAVID.

The full list of features associated with sensitivity (genes with negative effect size) was used as input in the functional analysis tool at http://david.abcc.ncifcrf.gov/[6]. GO terms enrichment was performed using the default settings of the DAVID tool V6.7.

Chronic lymphocytic leukemia. PBMCs freshly isolated from peripheral blood are seeded in 24 well plates at a volume of 0.5ml and a concentration of 2x10^6 cells/ml in IMDM medium supplemented with 10% Human AB serum, Transferrin, Insulin and Pen-Strep. Compounds and controls are added to respective wells, and incubated approximately 24 hrs at 37C, 5% CO2. Apoptosis is measured using the Annexin V: FITC Apoptosis Detection Kit I (BD Biosciences). Percentage cell death includes Annexin V and PI single and double positive cell populations, and is normalized to DMSO control wells using the calculation: [(% Dead Drug - % Dead DMSO)/(100% - % Dead DMSO)]*100.

Fluorescence-Activated Cell Sorting Analysis (FACS). For cell cycle analysis, cells were treated with inhibitors for indicated time periods. Cells were collected by centrifugation, washed once in ice-cold phosphate-buffered saline (PBS), and fixed overnight at −20 ^{o}C with 80% ethanol in PBS. Cells were washed three times with PBS. Finally, cells were resuspended in PBS containing 0.1% Triton X-100, 25 μg· ml^{-1}

propidium iodide (PI, Molecular Probes cat# P1304MP), and 0.2 $mg \cdot ml^{-1}$ RNase A (Sigma cat# R4642) and incubated for 45 minutes at 37 ^{o}C. For discrimination of apoptotic vs. non-apoptotic cells by Annexin V/ PI staining, cells were treated with compound for indicated periods of time. Cells were collected by centrifugation, washed once with PBS, and processed according to manufacturers protocol (Invitrogen cat# V13242). All FACS samples (cell cycle distribution or PI/ Annexin V double staining) were analyzed on a BD LSR (BD Biosciences) instrument and processed on FlowJo (Treestar). All PI/ Annexin V experiments were performed in biological triplicate. Error bars are +/- SD.

KOPTK1 T-ALL human xenograft mouse efficacy study. Thirty-two NOD-SCID-IL2Rcγ^{null} (NSG) 9-week old female mice (The Jackson Laboratory, Bar Harbor, ME) were injected intravenously with $2x10^6$ KOPTK-1 cells expressing luciferase. Leukemia burden was established by bioluminescence imaging (BLI) using an IVIS Spectrum system (PerkinElmer Health Sciences, Shelton, CT) beginning one week following cell injection. At this time, mice were divided into treatment groups based on mean BLI as follows: THZ1 10mg/kg qD, THZ1 10mg/kg BID, and vehicle (10% DMSO in D5W) BID (n=10 for all groups). Two mice were excluded, one with the highest and one with the lowest BLI. All treatments were administered via IV injection in the lateral tail vein in a volume of 3.3µL/g (non-blinded). Mice were imaged and weighed every 3-5 days. Mice were treated for four weeks and on the final day mice were imaged, dosed and sacrificed approximately 5-6 hrs post dose. Upon sacrifice, blood was collected via cardiac puncture in EDTA tubes; a portion (~300 uL) was processed for plasma. Liver

and spleen tissues were collected from each mouse with half of each sample flash frozen and half of each sample fixed. Blood plasma and liver samples were processed for pharmacokinetics analysis of THZ1. Spleen tissues were homogenized and lysed and processed for pharmacodynamics analysis of THZ1 target engagement. All animal studies were performed using protocols approved by the Institutional Care and Use Committees of the Dana-Farber Cancer Institute and the Whitehead Institute for Biomedical Research.

Pharmacokinetics study of THZ1 from KOPTK1 T-ALL human xenograft mouse efficacy study samples. The concentration of THZ1 was determined in plasma and liver samples obtained at the end of the efficacy study. Plasma was generated using standard centrifugation techniques and liver samples were snap frozen and stored at -80^{O}C until analysis. On the day of analysis, plasma and liver samples were thawed on ice, mixed with acetonitrile (1:5 v:v or 1:5 w:v, respectively), sonicated with a probe tip sonicator, and centrifuged through a Millipore Multiscreen Solvinter 0.45 micron low binding PTFE hydrophilic filter plate prior to analysis by LC-MS/MS. Concentrations were determined using an ABSciex 5500 mass spectrometer. THZ1 was detected using a mass transition of 566.4 to 186.1. Plasma drug levels were quantitated using standards made in blank mouse plasma and liver levels against standards made in blank mouse liver homogenate.

Patient-derived xenograft (PDX) treatment with THZ1. Patient-derived xenograft cells were treated with THZ1 for 3 hrs in vitro followed by compound washout (3 washes with DPBS). An aliquot of input cells was then counted by flow cytometry using a known

quantity of flow cytometry calibration beads (data not shown; Molecular Probes). The remaining cells were plated onto MS5-DL1 feeder cells in the presence of serum-free media[7]. 72 hrs later, cultures were harvested by vigorous pipetting with trypsin, filtered through nylon mesh to deplete feeders, and counted by flow cytometry using a known quantity of flow cytometry calibration beads. Based on an estimated average of 1.6-fold expansion with standard deviation of 0.7 and estimated minimum average of 0.4 following treatment with compound (i.e. at least 75% cell death), the test would be 80% powered for a p=0.05 difference with 3 samples.

RNA Extraction and Synthetic RNA Spike-In. Total RNA and sample preparation was performed as previously described[8]. Briefly, Jurkat cells were incubated in media containing THZ1, THZ1-R, Flavopiridol, or Actinomycin D at the indicated concentrations or with DMSO for the specified duration of time. Cell numbers were determined by manually counting cells using C-Chip disposable hemocytometers (Digital Bio, DHC-N01) prior to lysis and RNA extraction. Biological duplicates (equivalent to 5 million cells per replicate) were subsequently collected and homogenized in 1 ml of TRIzol Reagent (Life Technologies, 15596-026), purified using the mirVANA miRNA isolation kit (Ambion, AM1560) following the manufacturer's instructions and re-suspended in 50 μL nuclease-free water (Ambion, AM9938). Total RNA was spiked-in with ERCC RNA Spike-In Mix (Ambion, 4456740), treated with DNA-freeTM DNase I (Ambion, AM1906) and analyzed on Agilent 2100 Bioanalyzer for integrity. RNA with the RNA Integrity Number (RIN) above 9.8 was hybridized to GeneChip PrimeView Human Gene Expression Arrays (Affymetrix).

cDNA Preparation and TaqMan Expression Analysis. RNA utilized for RT-PCR was extracted as outlined above. One microgram of purified RNA was reverse transcribed using Superscript III First-Strand (Invitrogen, cat# 18080-051) with oligo dT primers to prime first-strand synthesis according to the manufacturer's protocol. qPCR was carried out on the 7000 ABI Detection System using the following Taqman probes according to the manufacturer's protocol (Applied Biosystems). All experiments shown were performed in biological triplicate. Each individual biological sample was qPCR-amplified in technical triplicate. Error bars are +/- SD. Expression was normalized to ACTB, and fold change in expression was calculated relative to the indicated conditions. Error bars represent standard deviation. The following Taqman probes from Life Technologies were used for qPCR-based gene expression analysis: RUNX1 - Hs00231079_m1; TAL1 - Hs01097987_m1; GATA3 - Hs00231122_m1; MYB - Hs00920556_m1; MYC - Hs00153408_m1; BCL2 - Hs00608023_m1; MCL1 - Hs01050896_m1; ACTB - Hs01060665_g1; ERG - Hs01554629_m1; ARID5B - Hs01382781_m1; LYL1 - Hs01089802_g1; BCL6 - Hs00153368_m1; BCL11A - Hs01093197_m1; LMO2 - Hs00153473_m1 ; XBP1 - Hs00231936_m1; PRDM1 - Hs00153357_m1; BCL2L1 - Hs00236329_m1.

Microarray Sample Preparation and Analysis

For microarray analysis, 100 ng of total RNA containing ERCC RNA Spike-In Mix (see above) was used to prepare biotinylated aRNA (cRNA) according to the manufacturer's protocol (30 IVT Express Kit, Affymetrix 901228). Briefly, total RNA undergoes T7

oligo(dT)-primed reverse transcription to synthesize first-strand cDNA containing a T7 promoter sequence. This cDNA is then converted into a double-stranded DNA template for transcription using DNA Polymerase and RNase H to simultaneously degrade the RNA and synthesize second strand cDNA. In vitro transcription synthesizes aRNA and incorporates a biotin-conjugated nucleotide. The aRNA is then purified to remove unincorporated NTPs, salts, enzymes, and inorganic phosphate. Fragmentation of the biotin-labeled aRNA prepares the sample for hybridization onto GeneChip 3' expression arrays. Samples were prepared for hybridization using 10 μg of biotinylated aRNA in a 1X hybridization cocktail according the Affymetrix hybridization manual. Additional hybridization cocktail components were provided in the Affymetrix GeneChip Hybridization, Wash and Stain Kit. GeneChip arrays (Human PrimeView, Affymetrix 901837) were hybridized in a GeneChip Hybridization Oven at 45°C for 16 hrs at 60 RPM. Washing was done using a GeneChip Fluidics Station 450 according to the manufacturer's instructions, using the buffers provided in the Affymetrix GeneChip Hybridization, Wash and Stain Kit. Images were extracted with Affymetrix GeneChip Command Console (AGCC), and analyzed using GeneChip Expression Console. A Primeview CDF that included probe information for the ERCC controls (GPL16043), provided by Affymetrix, was used to generate .CEL files. We processed the CEL files using standard tools available within the affy package in R. The CEL files were processed with the expresso command to convert the raw probe intensities to probeset expression values. The parameters of the expresso command were set to generate Affymetrix MAS5-normalized probeset values. We used a *loess* regression to re-normalize these MAS5 normalized probeset values, using only the spike-in probesets to

fit the *loess*. The affy package provides a function, *loess* normalize, which will perform *loess* regression on a matrix of values (defined using the parameter mat) and allows for the user to specify which subset of data to use when fitting the *loess* (defined using the parameter subset, see the affy package documentation for further details). For this application the parameters mat and subset were set as the MAS5-normalized values and the row-indices of the ERCC control probesets, respectively. The default settings for all other parameters were used. The result of this was a matrix of expression values normalized to the control ERCC probes. The log2 probeset values from the duplicates were averaged together and the Log2 fold change comparing the log2 control to the log2 inhibitor- treated samples are shown.

PrimeView microarrays were submitted in duplicate in the following batches:

- Jurkat - 4h DMSO (2 replicates), 4h 50nM THZ1 (2 replicates)
- Jurkat - 6h DMSO (2 replicates), 6h 250nM THZ1 (2 replicates), 6h 250nM THZ1-R (2 replicates), 12h DMSO (2 replicates), 12h 250nM THZ1 (2 replicates)
- Jurkat - 4h DMSO for actinomycin D and flavopiridol comparison (2 replicates), 4h 1uM actinomycin D (1 replicate), 4h 1uM flavopiridol (1 replicate), 4h DMSO for THZ1 comparison (2 replicates), 4h 250 nM THZ1
- Jurkat - 0h DMSO (2 replicates), 0.25h 1uM flavopiridol (2 replicates), 0.5 1uM flavopiridol (2 replicates), 1h 1uM flavopiridol (2 replicates), 2h 1uM flavopiridol (2 replicates), 4h 1uM flavopiridol (2 replicates), 6h 1uM flavopiridol (2 replicates), 8h 1uM flavopiridol (2 replicates), 12h 1uM flavopiridol (2 replicates)

- HeLa – wild-type Dox- 0nM THZ1 (2 replicates), wild-type Dox- 100nM THZ1 (2 replicates), wild-type Dox- 500nM THZ1 (2 replicates), wild-type Dox+ 0nM THZ1 (2 replicates), wild-type Dox+ 100nM (2 replicates), wild-type Dox+ 500nM THZ1 (2 replicates), C312S Dox- 0nM THZ1 (2 replicates), C312S Dox- 100nM THZ1 (1 replicate), C312S Dox- 500nM THZ1 (2 replicates), C312S Dox+ 0nM THZ1 (2 replicates), C312S Dox+ 100nM THZ1 (2 replicates), C312S Dox+ 500nM (2 replicates)

Batches of microarray experiments were normalized within themselves to ERCC spike-in probes as described in[8]. Briefly, probes were combined into probesets using *expresso* with mas5 normalization. Probeset expression profiles were normalized using *loess* from the *affy* R package to equilibrate ERCC spike-in probes across microarrays in a given batch. Where possible, log2 values of biological replicates were averaged. Fold-changes were taken by subtracting average log2 DMSO signal from average log2 treatment signal. Expressed genes were those with log2(expression) > log2(100) in the corresponding DMSO sample.

Gene Set Enrichment Analysis (Figure 4f).

Gene Set Enrichment Analysis[9] was performed using a gene list pre-ranked by fold-change upon Runx1 knockdown from[10]. The 500 genes expressed in 4h DMSO from the same microarray batch as 4h 50nM THZ1 and most significantly downregulated by 4h 50nM THZ1 treatment was used as the gene set.

Chromatin Immunoprecipitation. Cells were crosslinked for 10 min at room

temperature by the addition of one-tenth of the volume of 11% formaldehyde solution (11% formaldehyde, 50mM HEPES pH 7.3, 100mM NaCl, 1mM EDTA pH 8.0, 0.5mM EGTA pH 8.0) to the growth media followed by 5 min quenching with 100 mM glycine. Cells were washed twice with PBS, then the supernatant was aspirated and the cell pellet was flash frozen in liquid nitrogen. Frozen crosslinked cells were stored at -80°C. 50 µL of Dynal magnetic beads (Sigma) were blocked with 0.5% BSA (w/v) in PBS. Magnetic beads were bound with 5 µg of the indicated antibody. For RNA polymerase II occupied genomic regions, we performed ChIP-Seq experiments using a SantaCruz Biotechnology (sc-899, lot K0111) antibody. The affinity purified antibody was raised in rabbit against an epitope mapping to the N-terminus of murine RBP1, the largest subunit of RNA Pol II. For CDK7 occupied genomic regions, we performed ChIP-Seq experiments using a Bethyl Laboratories (A300-405A-1) antibody. The affinity-purified antibody was raised in rabbit against an epitope corresponding to amino acids 300-346 of human CDK7. For H3K27Ac occupied genomic regions, we performed ChIP-Seq experiments using an Abcam (AB4729A) antibody. The affinity-purified antibody was raised in rabbit against an epitope corresponding to amino acids 1-100 of human Histone H3 that is acetylated at K27. For Jurkat, crosslinked cells were lysed with lysis buffer 1 (50mM HEPES pH 7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP-40, and 0.25% Triton X-100), pelleted and resuspended in lysis buffer 2 (10 mM TrisHCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA). The subsequent pellet was resuspended in and sonicated in sonication buffer (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA pH 8.0, 1 mM EGTA, 0.1% Na-deoxycholate, 0.1% SDS, and 1% Triton X-100). Cells were sonicated for 10 cycles at 30 s each on ice (18-21 W) with 60 s on ice between cycles. Sonicated

lysates were cleared and incubated overnight at 4ºC with magnetic beads bound with antibody to enrich for DNA fragments bound by the indicated factor. Beads were washed two times with sonication buffer, one time with sonication buffer with 500 mM NaCl, one time with LiCl wash buffer (10 mM TrisHCl pH 8.0, 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate) and one time with TE. DNA was eluted in elution buffer (50 mM TrisHCl pH 8.0, 10 mM EDTA, 1% SDS). Cross-links were reversed overnight. RNA and protein were digested using RNase A and Proteinase K, respectively and DNA was purified with phenol chloroform extraction and ethanol precipitation.

ChIP-Seq Analysis (Figure 4, Supplementary Figure 10).

Illumina sequencing libraries generated and data was processed according to[11]. In brief, libraries were generated for ChIP samples following the Illumina TruSeq™ DNA Sample Preparation v2 kit protocol with minor changes. All ChIP-Seq data sets were aligned using bowtie 0.12.9[12] to build NCBI36/hg18 of the human genome with parameters -e 70, -k 1, -m 1, -n 2, –best, –sam. Wiggle files for gene tracks were created using Macs 1.4.2[13] with options –w –S –space=50 to count reads in 50bp bins, were divided by the number of treatment reads to normalize to mapped-reads-per-million, and were displayed in the UCSC genome browser.

Metagene Analysis (Figure 4b).

Metagene analyses were created for those Refseq transcripts in the hg18 genome that are expressed at 6h by microarray (mean log2(DMSO expression) > log2(100)). These genes were divided into 50 intragenic bins. 2kb upstream and downstream of each gene were

each divided into 50 bins. Reads were artificially extended 200bp and their reads-per-million normalized density was calculated in each bin. The mean density per bin is displayed.

Super-Enhancer Identification (Supplementary Figure 10b and Supplementary Table 8)

Super-enhancers were identified as described in[14,15] with a stitching distance of 12.5kb. Briefly, Macs peaks (-p 1e-9, --keep-dup=auto, with input control) of H3K27ac were used to identify constituent enhancers. H3K27ac signal (less input control) was used to rank enhancers by their enrichment. In Jurkat 631 super-enhancers were separated from typical enhancers by finding the inflection point of the curve, respectively. Super-enhancers were assigned to the gene whose TSS falls nearest to the center of the super-enhancer. Jurkat super-enhancer #6 was originally assigned to C21orf96, a non-coding RNA located in the 1st intron of RUNX1. However, we assigned it to RUNX1 based on 1) visual inspection of H3K27Ac and RNAPII signal at the RUNX1 and C21orf96 genes and 2) independent literature observations that demonstrated physical interaction between this enhancer region and both RUNX1 promoters using 3C mapping[16,17].

Supplementary References:

1 Sarbassov, D. D., Guertin, D. A., Ali, S. M. & Sabatini, D. M. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307**, 1098-1101, doi:10.1126/science.1106148 (2005).

2 Lolli, G., Lowe, E. D., Brown, N. R. & Johnson, L. N. The crystal structure of human CDK7 and its protein recognition properties. *Structure* **12**, 2067-2079, doi:10.1016/j.str.2004.08.013 (2004).

3 Ficarro, S. B. *et al.* Improved electrospray ionization efficiency compensates for diminished chromatographic resolution and enables proteomics analysis of tyrosine signaling in embryonic stem cells. *Analytical chemistry* **81**, 3440-3447, doi:10.1021/ac802720e (2009).

4 Parikh, J. R. *et al.* multiplierz: an extensible API based desktop environment for proteomics data analysis. *BMC bioinformatics* **10**, 364, doi:10.1186/1471-2105-10-364 (2009).

5 Garnett, M. J. *et al.* Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature* **483**, 570-575, doi:10.1038/nature11005 (2012).

6 Huang da, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research* **37**, 1-13, doi:10.1093/nar/gkn923 (2009).

7 Yost, A. J. *et al.* Defined, serum-free conditions for in vitro culture of primary human T-ALL blasts. *Leukemia* **27**, 1437-1440, doi:10.1038/leu.2012.337 (2013).

8 Loven, J. *et al.* Revisiting global gene expression analysis. *Cell* **151**, 476-482, doi:10.1016/j.cell.2012.10.012 (2012).

9 Mootha, V. K. *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature genetics* **34**, 267-273, doi:10.1038/ng1180 (2003).

10 Sanda, T. *et al.* Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. *Cancer cell* **22**, 209-221, doi:10.1016/j.ccr.2012.06.007 (2012).

11 Lin, C. Y. *et al.* Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **151**, 56-67, doi:10.1016/j.cell.2012.08.026 (2012).

12 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).

13 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).

14 Loven, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320-334, doi:10.1016/j.cell.2013.03.036 (2013).

15 Whyte, W. A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307-319, doi:10.1016/j.cell.2013.03.035 (2013).

16 Markova, E. N., Kantidze, O. L. & Razin, S. V. Transcriptional regulation and spatial organisation of the human AML1/RUNX1 gene. *Journal of cellular biochemistry* **112**, 1997-2005, doi:10.1002/jcb.23117 (2011).

17 Nottingham, W. T. *et al.* Runx1-mediated hematopoietic stem-cell emergence is controlled by a Gata/Ets/SCL-regulated enhancer. *Blood* **110**, 4188-4197, doi:10.1182/blood-2007-07-100883 (2007).